

Experimental economics for web mining

Rustam Tagiev
Alumni of
TU Bergakademie Freiberg
yepkio@mail.ru

Dmitry I. Ignatov
National Research University
Higher School of Economics
dignatov@hse.ru

Fadi Amroush
Granada Lab of
Behavioral Economics (GLOBE)
Granada, Spain
fadyamr@gmail.com

Abstract—This paper offers a step towards research infrastructure, which makes data from experimental economics efficiently usable for analysis of web data. We believe that regularities of human behavior found in experimental data also emerge in real world web data. A format for data from experiments is suggested, which enables its publication as open data. Once standardized datasets of experiments are available on-line, web mining can take advantages from this data. Further, the questions about the order of causalities arisen from web data analysis can inspire new experiment setups.

I. INTRODUCTION

From the view point of web miners or rather data scientists dealing with human behavior, the sizes of datasets originated from web predominate the ones from experimental economics by orders of magnitude [1]. Nevertheless, analyzing data from experimental economics has the same importance for understanding human psychology as studying *E. coli* for understanding human physiology. It is reasonable to assume that human behavior regularities found in experimental data also emerge in web data and vice versa. Data from experimental economics has the advantage of originating from simple and controlled human interactions.

The reason for economists to conduct experiments was the deficiency of the pure theory. It is not only the global financial crisis of the recent years [2], which made economists reconsider the path economics as a discipline should take. Since decades, it became obvious that classical theories fail in real world [3]. Paul Krugman described the current situation in economics as: "... the central cause of the professions failure was the desire for an all-encompassing, intellectually elegant approach that also gave economists a chance to show off their mathematical prowess. Unfortunately, this romanticized and sanitized vision of the economics led most economists to ignore all the things that can go wrong. They turned a blind eye to the limitations of human rationality".

Experimental economics gained its importance as a promising solution of this problem. Experiments on human subjects have shifted economists' point of view closer to the psychologists' one – people are no more considered to be rational monetary payoff maximizers alias *homines oeconomici* [4]. This makes economic research more useful for web mining, since web users do not always interact in monetized or monetizable domains. Laboratory and field experiments are carried out using human subjects in order to improve theoretical knowledge about human behavior in

interactions. The body of knowledge or rather the discipline arising from this process is called behavioral economics.

Today, behavioral economists simplify real world cases to an experiment setup, make an assumption about the behavior model, conduct experiments, fit their model's parameters and come back with an explanation of the real world human behavior. Unfortunately, many presumed models from economics might be conceptually not appropriate to fit human behavior. For instance, linear models are commonly used on nominal data [5], [6], whereby nominal value sets are arbitrarily projected into real numbers. Otherwise, non-deterministic models as QRE [7] are used to explain human behavior, although human subjects fail at generating truly random sequences and at dealing with probabilities [8], [9]. It is recommended to reanalyze data in some cases instead of trusting the published results. Therefore, a facility for publishing original experimental data has to be created. In fact, closing the gap between behavioral economics and data science remains a challenge for future.

As for behavioral economics, as for web mining and as for any scientific discipline, real world applications have to be at least the long-term results. Behavioral economists and web miners are confronted with the same problem – they have to predict human behavior and to develop policies to steer it the desired way. Starting from a real world question, behavioral economists and web miners branch into two different and mutually completing approaches. Behavioral economists can find true causalities, as long these causalities match into simple model spaces. Whereby, web miners can generate complicated models, but are not able to decide the order of causalities. Once the expert knowledge from behavioral economists, formatted experimental data, advanced machine learning algorithms and web data come together, more reliable solutions can be provided to real world problems.

II. RELATED WORK

Regarding open experimental data, there was the by now defunct website ExLab (exlab.bus.ucf.edu) launched by the department of Economics of University of Central Florida. This website included only data from non-interactive experiments. Further, the data was not uniformly formatted.

Market leaders already push forward into the area between behavioral economics and data science. Facebook hired economists for studying user and advertiser behavior,

TABLE I
5 INDEPENDENT DATA SETS.

Dataset	Number of human decisions
Roshambo [18]	4000
Colored Trails (Responder) [19]	371
Product Selection [20]	145
Social Learning [21]	14040
Gift-Exchange-Game [6]	746

economic networks, incentives, externalities, and decision making under risk and uncertainty [10]. Microsoft founded MSR-NYC [11], where researchers develop technologies in the intersection of social science, both computational and behavioral, computational economics and prediction markets, machine learning, as well as information retrieval. Google’s chief economist openly writes in his paper [12]: “I believe that [manipulating and analyzing big data] have a lot to offer and should be more widely known and used by economists. In fact, my standard advice to graduate students these days is ‘go to the computer science department and take a class in machine learning’.”

In the academic world, the international workshop series “Experimental Economics and Machine Learning” (EEML) started in 2012 [13], [14], [15]. EEML seeks to fill the gap between two scientific communities of Experimental Economics and AI & Data Mining. The conference “Social media and behavioral economics” took place in 2013 [16], where data scientists and economists from universities and industry participated. Yale University has a chair researching in machine learning, behavioral economics, and finance [17].

III. DATA FORMAT FOR EXPERIMENTAL ECONOMICS

The authors of this paper have the cumulated experience on analyzing 5 independent datasets from experimental economics [22], [20], [23], [24]. Tab.I shows the number of single human decision in every of the datasets. As a single human decision, we define one no further indivisible turn executed by one human. For every turn, there is some information visible to the decision taking subject. In a subsequent decision, this information also includes own previous decisions and the information, which was visible earlier. Once a set of decision samples with equal structure of visible information is created, data mining can be conducted. Tab.II shows an example of a refined set of decision samples.

From our experience, the data refinery until this point made by an external data scientist takes more time as the actual data mining. Especially, determining information visible to a player is a time consuming process. Therefore, we suggest a format for the interchange between economists and data scientists, where all same structured decision samples are organized in separated sets as already exemplarily shown on Tab.II. We further call visible information as inputs. For instance, the “Colored Trails” dataset has only one set of decision samples with ‘own payoff’ and ‘partner’s payoff’ as

inputs.

IV. COMMUNITY FORMATION

The issues expressed before should not reduce experimental economics to data creation, although the knowledge creation conducted solely by economists is not optimal. We rather suggest to relocate economists from the start to the end of the knowledge creation process. That means that the data is analyzed first by data scientists and then the regularities are corrected and interpreted by economists. Therefore, every research project has to include data scientists as well as economists in its participants’ list. A web resource offering uniformly formatted datasets should serve as a focal point for the crystallization of such research teams. The data license is an issue to be solved for such a facility.

V. CONCLUSION

We suggested a simple format for efficient interchange of experiment data between economists and data scientists, which should facilitate the cooperation between them to conduct research on real world problems concerning human behavior on web.

TABLE II
EXAMPLE FOR A REFINED SET OF DECISION SAMPLES. DAVID’S DECISION IS TO PREDICT.

World	Alice	Bob	Charly	David
sun	beach	beach	library	library
rain	library	cinema	library	library
cloudy	park	park	beach	beach
⋮	⋮	⋮	⋮	⋮

REFERENCES

- [1] C. F. Camerer, *Behavioral Game Theory*. Princeton University Press, 2003.
- [2] G. P. Maxton, *The end of progress : how modern economics has failed us*. John Wiley & Sons, 2011.
- [3] D. Ariely, *Predictably Irrational*. Harper, 2009.
- [4] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath, "In search of homo economicus: Behavioral experiments in 15 small-scale societies," *The American Economic Review*, vol. 91, no. 2, pp. 73–78, 2001.
- [5] C. R. Plott and V. L. Smith, Eds., *Handbook of Experimental Economics Results*. North-Holland, 2008.
- [6] S. Gächter, D. Nosenzo, and M. Sefton, "The impact of social comparisons on reciprocity," *The Scandinavian Journal of Economics*, vol. 114, no. 4, pp. 1346–1367, 2012.
- [7] M. Richard and T. Palfrey, "Quantal response equilibria for normal form games," *Games and Economic Behavior*, vol. 10, pp. 6–38, 1995.
- [8] W. A. Wagenaar, "Generating of random sequences by human subjects: A critical survey of the literature," *Psychological Bulletin*, vol. 77, pp. 65–72, 1972.
- [9] M. Oaksford and N. Chater, "The probabilistic approach to human reasoning," *Trends in Cognitive Sciences*, vol. 5, pp. 349–357, 2001.
- [10] M. Bailey, "What does an economist at facebook do?" [quora.com/What-does-an-economist-at-Facebook-do/](https://www.quora.com/What-does-an-economist-at-Facebook-do/), 2014.
- [11] Microsoft Inc., "Microsoft research new york city," research.microsoft.com/en-us/labs/newyork/, 2014.
- [12] H. R. Varian, "Big data: New tricks for econometrics," people.ischool.berkeley.edu/~hal/Papers/2013/ml.pdf, 2013.
- [13] R. Tagiew, D. I. Ignatov, A. A. Neznanov, and J. Poelmans, Eds., *First International Workshop on Experimental Economics and Machine Learning*. KU-Leuven, 2012, ceur-ws.org/Vol-870/, at ICFCFA.
- [14] R. Tagiew, D. I. Ignatov, and F. Amroush, Eds., *Second International Workshop on Experimental Economics and Machine Learning*. IEEE Computer Society, 2013, [dx.doi.org/10.1109/ICDMW.2013.178](https://doi.org/10.1109/ICDMW.2013.178), at ICDM.
- [15] R. Tagiew, D. I. Ignatov, F. Amroush, and B. Gutkin, "Third international workshop on experimental economics and machine learning," eeml.hse.ru/2014/, 2014.
- [16] C. Sunstein and J. Zittrain, "Social media and behavioral economics," today.law.harvard.edu/social-media-and-behavioral-economics-conference/, 2013.
- [17] "Philip R. Allen professor of economics," som.yale.edu/donald-j-brown, Yale school of management.
- [18] R. Tagiew, "Hypotheses about typical general human strategic behavior in a concrete case," in *AI*IA*. Springer, 2009, pp. 476–485.
- [19] D. Antos and A. Pfeffer, "Using reasoning patterns to help humans solve complex games," in *IJCAI*. AAAI Press, 2009, pp. 33–39.
- [20] F. Amroush, "Proposing a similarity measure in case based reasoning for products selection – an experimental evidence," in *ICAART*, 2012, pp. 499–502.
- [21] S. Choi, D. Gale, and S. Kariv, "Social learning in networks: a quantal response equilibrium analysis of experimental data," *Review of Economic Design*, vol. 16, no. 2, pp. 135–157, 2012.
- [22] R. Tagiew, "Mining determinism in human strategic behavior," in *EEML*. KU-Leuven, 2012, pp. 86–91.
- [23] R. Tagiew, D. Ignatov, and F. Amroush, "Social learning in networks: Extraction of deterministic rules," in *Experimental Economics and Machine Learning*, ser. ICDM Workshops. IEEE Computer Society, 2013.
- [24] R. Tagiew and D. I. Ignatov, "Reciprocity in gift-exchange-games," *CoRR*, vol. abs/1402.5593, 2014.